Graphics, Vision and Image Processing Journal, ISSN 1687-398X, Volume 16, Issue 3, ICGST LLC, Delaware, USA, Dec. 2016

www.icgst.com

# Spectral-Spatial Classification of Hyperspectral Image based on Oversampling and Multi-Feature Kernels

Rafika Ben Salem[1], Karim Saheb Ettabaa[2,3] and Mohamed Ali Hamdi[1]
[1]*Laboratory for Materials, Molecular and applications, National Institute of Applied Sciences and Technology, Tunis, Tunisia.*
bs_rafika@yahoo.fr, mohamedalihamdi@yahoo.fr
[2]*Laboratory for research in computer Arabized and integrated documentation, National School of Computer Sciences, Tunis, Tunisia.*
[3]*Laboratory ITI, Telecom Bretagne, Brest Iroise Technopole CS 81828*
karim.sahebettabaa@riadi.rnu.tn

## Abstract
Spectral-Spatial classification of hyperspectral image suffers from two problems: the existence of various feature extraction methods that complicate the choice of those applying and the availability of limited number of labeled training samples. To overcome these difficulties, this paper presents new spectral-spatial classification approach for remotely sensed hyperspectral image which integrate different spectral and spatial features via multi-feature kernels and process accurately with limited number of training samples. In fact, the proposed method introduces different methods to extract the spectral and the spatial features and exploits the oversampling based on interpolation techniques to generate new labeled samples. First, each pixel must be characterized by two spectral vectors computed according to the application of the principal components analyse and the independent components analyse and three spatial features calculated by using three methods: the average of neighbourhood pixels, the textural features and the extended multi-attribute profiles. Then an oversampling step is introduced to create new labeled samples used to train the classifier. Finally, a support vector machine (SVM) with multi-feature kernel is efficiently trained to generate the classification map. The proposed classification approach is experimentally evaluated using the AVIRIS Indian Pines data set, exhibiting higher performance when compared with the multi-feature classification without oversampling.

***Keywords:*** *Hyperspectral images, SVM, multi-feature kernels, interpolation techniques.*

***Nomenclature:***

| | |
|---|---|
| *AA* | Average Accuracy |
| *AP* | Attribute Profile |
| *ASM* | Angular Second Moment |
| *EAP* | Extended Attribute Profile |
| *ENT* | Entropy |
| *EMAP* | Extended Multi-Attribute Profile |
| *GLCM* | Gray Level Co-Occurrence Matrix |
| *ICA* | Independent Components Analyses |
| *k* | Kappa coefficient |
| *LM* | Local Mean |
| *MLR* | Multinomial Logistic Regression |
| *MP* | Morphological Profiles |
| *OA* | Overall Accuracy |
| *PCA* | Principal Components Analyses |
| *SVM* | Support Vector Machine |
| *Var* | Variance |

## 1. Introduction
Recent advances on remote sensing provide images with high spectral and spatial resolution. Hyperspectral image presents the captured scene in hundreds of narrow contiguous bands spanning the visible-to-infrared spectrum. For that, hyperspectral data are used in a diverse applications such as agriculture [1], astronomy [2], surveillance [3] and environmental sciences [4]. These application are based on the classification of each pixel in hyperspectral imagery. The objective of the classification is to assign each pixel to one of the classes, based on its spectral and spatial characteristics. Then, the exploitation of the highly informative spectral and spatial information of hyperspectral image pixels improves the accuracy of the classification. Nevertheless, the complexity and the high dimensionality of hyperspectral data complicate the classification, thus this technique is a challenging task.
In the last decades, many discriminative classification approaches have been developed. Among these, the SVM [5] and MLR [6]–[4] have demonstrated to be very powerful. In particular, SVM has shown good performances for classifying high-dimensional data [7]. For that various spectral-spatial classification methods based on SVM have been presented in literature.

Composite kernels [8], which combine spectral and spatial kernels have been used to assure an accurate classification. The uses of stocked vector that concatenate spectral and contextual information extracted by MP has shown a significant improvements [9]. Edge-preserving filters have been exploited to develop an accurate spectral-spatial classification outperforming the classification without filtering [10]. A multi-feature model aiming at constructing a SVM set combining multiple spectral and spatial features [11] has registered an accurate classification. A generalized composite kernels have been introduced to improve the performance of the classification [12]. Segmentation techniques have been investigated in [13]. A multi-feature kernels [14], which combine different type of kernels: kernel for each feature, have noted a relevant classification.

The literature review showed that all the studies emphasized the importance of the spatial information in the hyperspectral image classification. However, the availability of various spectral and spatial characterization methods (e.g. PCA [15], ICA [16], morphological features [17], wavelet-based texture [18], GLCM [19]) complicate the selection of the used methods.

Another difficulty has been discussed in the literature, the Hughes phenomenon referred to the high dimensionality of the hyperspectral data and the availability of a limited number of training samples. Then different solutions have been proposed to solve this problem. Among these we note: the application of features selection [20] and extraction methods to reduce the dimensionality of data and the uses of semi-supervised learning techniques to develop a semi-supervised classification approach based on the augmentation of the number of training samples from the set of unlabeled pixels. Synthetic data has been investigated in [21] to increase the set of labeled samples by oversampling, which generate new samples by means of interpolation techniques.

In this context, we propose a multi-feature spectral-spatial classification approach based on oversampling aims to solve the problem of the limited number of training samples and to overcome the difficulty of the choice of the adopted characterization methods. The proposed approach implements the following three main steps: 1) spectral and spatial characterization step that introduces different methods to extract the spectral and the spatial features, 2) oversampling which exploits interpolation techniques to create new labeled examples and 3) classification step based on the use of SVM with multi-feature kernels that combine different type of kernels: kernel for each feature.

The remainder of this paper is organized as follows. Section 2 describes the proposed approach. Section 3 reports classification results based on real hyperspectral data sets. Finally, Section 4 concludes with some remarks.

## 2. Proposed approach

The goal of the proposed approach is to have an accurate SVM classification dealing with these two problems:

- the existence of many features extraction methods and the difficulty of the choice of the suitable method,
- the limited number of training samples.

For that, we propose to apply an oversampling step presented in [21] to increase the number of labeled pixels and to use the multi-feature kernels to combine different attributes resulted from the application of various spectral and spatial feature extraction techniques.

The proposed approach implements the following three main steps: 1) spectral and spatial characterization step that introduces different methods to extract the spectral and the spatial features, 2) oversampling step that exploits interpolation techniques to create new labeled examples and 3) classification step based on the use of SVM with multi-feature kernels.

### *2.1. Spectral and spatial characterization*

The wealthy spectral and spatial information available in hyperspectral images allows for the possibility to distinguish between spectrally similar materials. Various methods have been widely used in the literature for spectral and spatial characterizing hyperspectral pixels. For the spectral characterization, authors usually used all the spectral information or dimensionality reduction techniques like PCA and ICA to extract the most informative data. For the spatial features extraction, different means have been adopted such as: features provided from the neighborhood of the pixel, attribute filters and textural features.

In this paper, we focus on the uses of PCA and ICA for the spectral characterization and the mean of neighborhood pixels, EMAP based on attribute filters and textural features for the spatial characterization.

- PCA: is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. Then, PCA aims to remove the correlation among the bands. In the process, the optimum linear combination of the original bands accounting for the variation of pixel values in an image is identified.
- ICA: is a popular approach to blind source separation, it has been investigated in the analyse of hyperspectral images to remove the dependence between bands.
- Average of neighbourhood pixels: this spatial characterization technique explains each pixel (p) in terms of its neighborhood ($p_k$) in a window (i*j) by calculating the average of their spectral information $X(p_k)$. It return $X_{avg}$ (Equation (1)).

$$X_{avg}(p)=\frac{1}{i*j}\sum_{k=1}^{i*j}X(p_k) \qquad (1)$$

- Textural features [22]: emphasize the texture structure of the graylevel image. They are local indexes computed by means of sliding windows of size P × Q. For hyperspectral image, these metrics can be found by adopting the panchromatic band, the first principal component or a discriminative band. Among these features we can note:

  - Local mean (LM): is computed on the graylevel values contained in the sliding window centered on the pixel $x_{ij}$. It return a local texture value $x_{ij}^{LM}$ (Equation (2)).

$$x_{ij}^{LM} = \frac{1}{PQ}\sum_{p,q\in w} x_{pq} \quad (2)$$

where w denotes the pixels contained in the window centered on $x_{ij}$.

- Variance (Var): It return a local texture value $x_{ij}^{Var}$ (Equation (3))

$$x_{ij}^{Var} = \frac{1}{PQ}\sum_{p,q\in w} (x_{pq} - x_{ij}^{ML})^2 \quad (3)$$

where w denotes the pixels contained in the window centered on $x_{ij}$ and $x_{ij}^{LM}$ the local mean of the considered pixel.

- Entropy (ENT): This measure compute the intensity of the texture in the considered image. It is based on the GLCM, that represents the relative occurrence frequency p(m, n) of two graylevel intensities m and n in the P × Q window at a given angular neighbourhood (Equation (4)).

$$x_{ij}^{ENT} = -\sum_{n_1}\sum_{n_2} p(n_1, n_2)\log p(n_1, n_2) \quad (4)$$

-Angular second moment (ASM): It indicates the local contrast. providing an accurate estimate on the degree of uniformity of the values of the GLCM (Equation (5)).

$$x_{ij}^{ASM} = \sum_{n_1}\sum_{n_2} p(n_1, n_2)^2 \quad (5)$$

- EMAP [23] is a profile that stacked the EAPs obtained using different type of attributes. The EAP is resulted by generating an AP (obtained by applying a sequence of attribute filters using various thresholds) on each of the first *p* principal components.

***2.2. Oversampling***

To increase the number of training samples, we implemented the oversampling algorithm (Algorithm 1) [21]. The goal of this algorithm is to generate from the *t* feature vectors $y_i$ of dimension *dim* presenting the set of training samples ($Y^{train}=(y_1, ...,y_t)$) *g* new feature vectors presenting the set of new training samples ($Y^{new}=(y^{new}_1, ..., y^{new}_g)$) by means of interpolation techniques. In fact, three interpolation methods have been used: linear interpolation, cubic spline interpolation and Lagrange interpolation.

**Algorithm 1:** *Oversampling*

***READ*** $Y^{train}$
    ***FOR*** *each row of the matrix* $Y^{train}$ *i (For i=1 to dim)*
        *Present each value of the row i by a point.*
        *Compute the interpolation function.*
        *Generate new samples abscissas.*
        *Compute new training samples according to the evaluation of the interpolation function f in new abscissas.*
        *Save new values in* $Y^{new}$.
    ***ENDFOR***
***PRINT*** $Y^{new}$

Figure 1 shows the flowchart of the oversampling step.

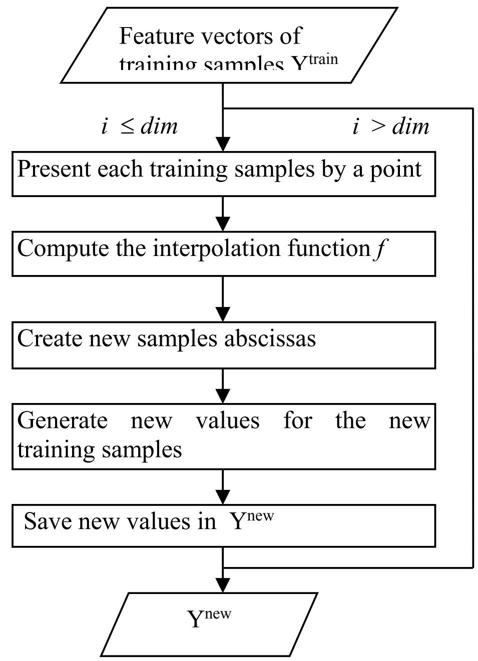



Figure 1: Flowchart of oversampling

***3.3. Classification via SVM with multi-feature kernels***

SVMs have been widely adopted in the classification of hyperspectral images due to their high performance registered on the process of data with high dimensionality. SVM [24] is a kernel based classifier consisting in projecting data in a higher dimension space by means of non-linear mapping function Φ and aiming at finding the optimal separator hyperplan by margin maximization. SVM has been proposed first for binary classification, after it has been introduced to solve multi-class classification.

In order to improve the classification performance achieved by using the spectral information alone, various spectral-spatial classification approaches incorporating the spatial information in addition to the spectral information have been proposed in the literature. In particular, the uses of SVM with kernels that combining different type of kernels like composite kernels [8] and multi-feature kernels [14] has shown high performance in term of accuracy.

In this paper, each pixel must be characterized by two spectral vectors $x^{PCA}$ and $x^{ICA}$ resulted respectively from the application of PCA et ICA and tree spatial vectors $x^{neighborhood}$, $x^{Texture}$ and $x^{EMAP}$ computed respectively after the implementation of the mean of neighborhood pixels, textural features and EMAP. For that, we implemented tree different multi-feature kernels which combining these different spectral and spatial attributes:

- Kernel 1 (Equation (6)):

$$K^{PCA+ICA+Neigh+Text+EMAP}(x_i, x_j) = K_{PCA}(x_i^{PCA}, x_j^{PCA}) + K_{ICA}(x_i^{ICA}, x_j^{ICA}) + K_{Neigh}(x_i^{Neigh}, x_j^{Neigh}) + K_{Text}(x_i^{Text}, x_j^{Text}) + K_{EMAP}(x_i^{EMAP}, x_j^{EMAP}) \quad (6)$$

- Kernel 2 (Equation (7)):

$$\mu K^{PCA+ICA+Neigh+Text+EMAP}(x_i, x_j) = \mu \times \left[ K_{PCA}(x_i^{PCA}, x_j^{PCA}) + K_{ICA}(x_i^{ICA}, x_j^{ICA}) \right] + (1-\mu) \times \left[ K_{Neigh}(x_i^{Neigh}, x_j^{Neigh}) + K_{Text}(x_i^{Text}, x_j^{Text}) + K_{EMAP}(x_i^{EMAP}, x_j^{EMAP}) \right] \quad (7)$$

with $0 < \mu < 1$.

- Kernel 3 (Equation (8)):

$$\alpha K^{PCA+ICA+Neigh+Text+EMAP}(x_i, x_j) = \alpha_1 \times K_{PCA}(x_i^{PCA}, x_j^{PCA}) + \alpha_2 \times K_{ICA}(x_i^{ICA}, x_j^{ICA}) + \alpha_3 \times K_{Neigh}(x_i^{Neigh}, x_j^{Neigh}) + \alpha_4 \times K_{Text}(x_i^{Text}, x_j^{Text}) + \alpha_5 \times K_{EMAP}(x_i^{EMAP}, x_j^{EMAP}) \quad (8)$$

where $\sum_{i=1}^{5} \alpha_i = 1$.

To summarize the description of our proposed classification method, Algorithm 2 provides a pseudocode for our newly developed spectral spatial classification algorithm based on a SVM classifier with multi-feature kernels and oversampling.

**Algorithm 2:** SVM_oversampling

**READ** $Y = (y_1, ..., y_n)$ // Pixels of the hyperspectral image.
**READ** T //Set of training samples.
*// Spectral characterization*
$Y^{PCA}$ = *PCA(Y) //Compute the PCA of each pixels.*
$T^{PCA} = [T_1^{PCA}, ..., T_c^{PCA}]$ *// compute PCA of labeled pixels.*
$Y^{ICA}$ = *ICA(Y) //Compute the ICA of each pixels.*
$T^{ICA} = [T_1^{ICA}, ..., T_c^{ICA}]$ *//compute ICA of learning pixels.*
*// Spatial characterization*
$Y^{Neigh}$ = *Neigh(Y) //Calculate for each pixel the average of neighborhood pixels.*
$T^{Neigh} = [T_1^{Neigh}, ..., T_c^{Neigh}]$ *//Average of neighborhood pixels of training samples.*
$Y^{EMAP}$ = *EMAP(Y) //Compute the EMAP of each pixels.*
$T^{EMAP} = [T_1^{EMAP}, ..., T_c^{EMAP}]$;
$Y^{Text}$ = *Text(Y) //Compute the textural features the of each pixels.*
$T^{Text} = [T_1^{Text}, ..., T_c^{Text}]$ //

**FOR** each class i
$Y_{i\,new}^{PCA}$ = *oversampling*($T_i^{PCA}$)
$Y_{i\,new}^{ICA}$ = *oversampling*($T_i^{ICA}$)
$Y_{i\,new}^{Neigh}$ = *oversampling*($T_i^{Neigh}$)
$Y_{i\,new}^{EMAP}$ = *oversampling*($T_i^{EMAP}$)
$Y_{i\,new}^{Text}$ = *oversampling*($T_i^{Text}$)
// Spectral and spatial features of training samples after oversampling.
$T_i^{PCA} = [T_i^{PCA}, Y_{i\,new}^{PCA}]$
$T_i^{ICA} = [T_i^{ICA}, Y_{i\,new}^{ICA}]$
$T_i^{Neigh} = [T_i^{Neigh}, Y_{i\,new}^{Neigh}]$
$T_i^{EMAP} = [T_i^{EMAP}, Y_{i\,new}^{EMAP}]$
$T_i^{Text} = [T_i^{Text}, Y_{i\,new}^{Text}]$
*// Features of training samples in all the classes.*
$T_{new}^{PCA} = [T^{PCA}, T_i^{PCA}]$
$T_{new}^{ICA} = [T^{ICA}, T_i^{ICA}]$
$T_{new}^{Neigh} = [T^{Neigh}, T_i^{Neigh}]$
$T_{new}^{EMAP} = [T^{EMAP}, T_i^{EMAP}]$
$T_{new}^{Text} = [T^{Text}, T_i^{Text}]$
**ENDFOR**
*L= SVM_Classification(* $Y^{PCA}, Y^{ICA}, Y^{Neigh}, Y^{EMAP}, Y^{Text}, T_{new}^{PCA}, T_{new}^{ICA}, T_{new}^{Neigh}, T_{new}^{EMAP}, T_{new}^{Text}$ *) // SVM classification with multi-feature kernels.*
*PRINT L // Labels of each pixel.*

Figure 2 illustrate the flowchart of algorithm 2.
The architecture of the proposed approach is illustrated in figure 3.

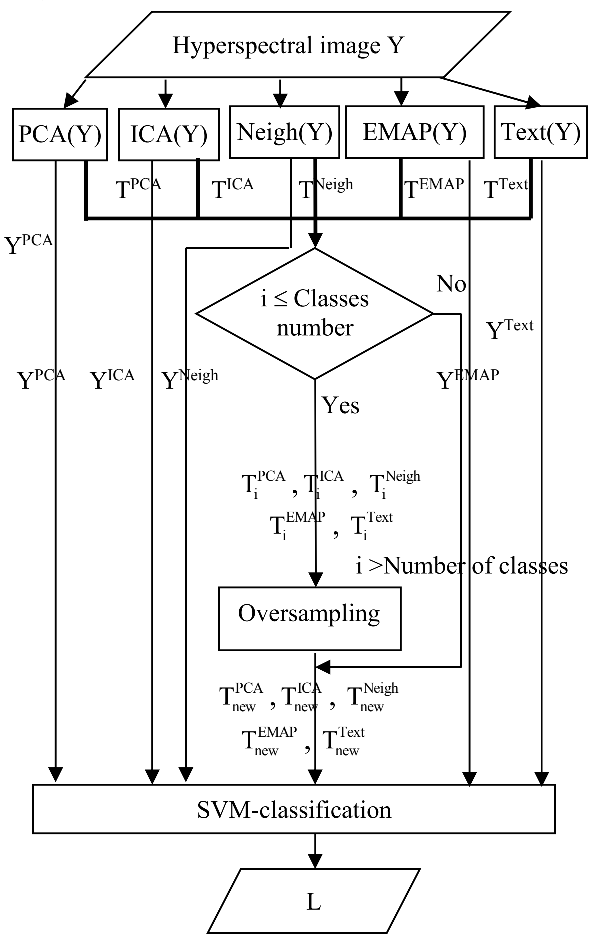


Figure 2: The flowchart of algorithm 2.

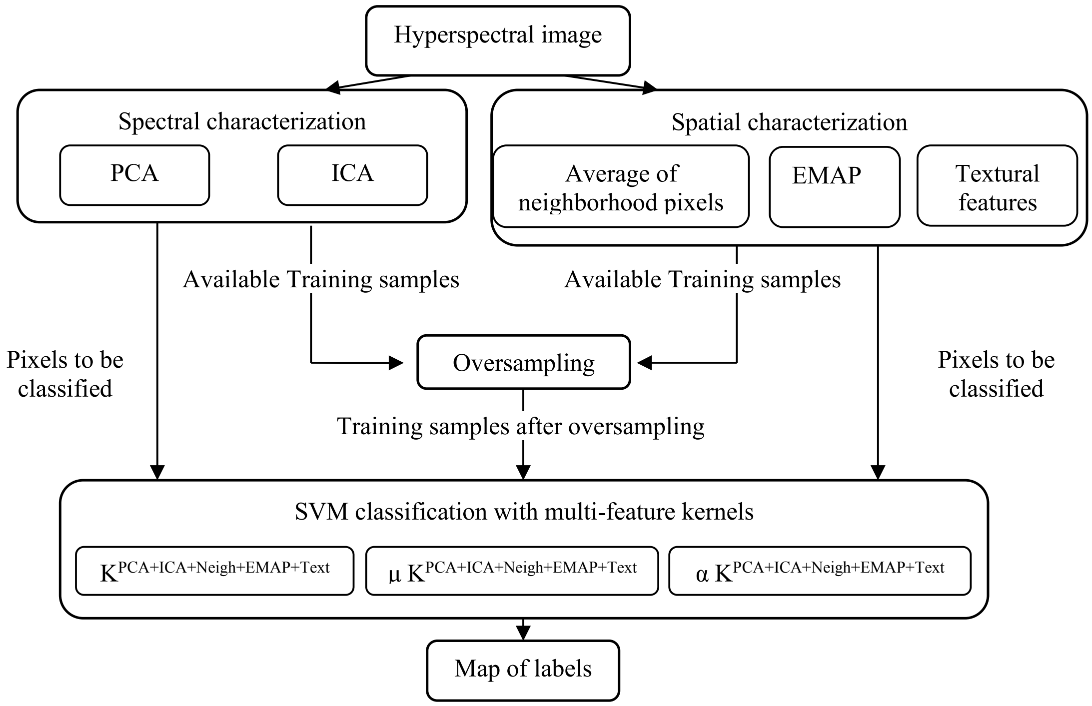



Figure 3: Architecture of the proposed approach.

## 3. Experimental results

To evaluate the performance of the proposed approach, we classified the widely used hyperspectral image " Indian Pines". It contains 145 × 145 pixels and 200 spectral bands. The ground-truth data contains 16 classes and a total of 10366 labeled pixels. The classification of this image is a challenging task refer to the significant presence of classes with similar spectral signatures and also because of the unbalanced number of available labeled pixels per class.

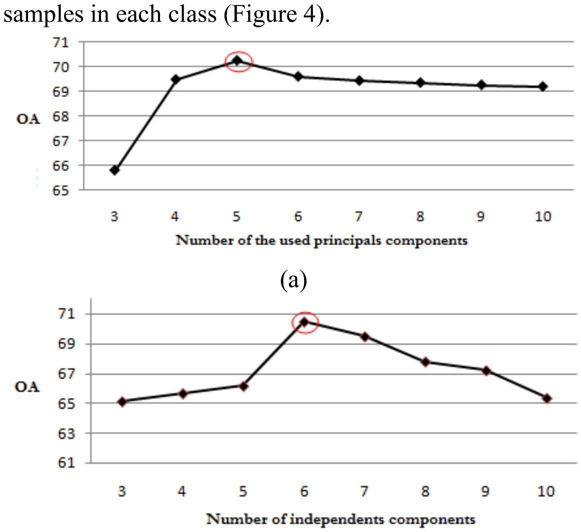
In all these experiments, we will use the classification accuracy (OA) and kappa coefficient (Kappa) as a references to evaluate the performance of the proposed classification approach.

For SVM classification, we implemented the most used multi-class classification strategy "one against all" and we used RBF and polynomial kernels for the spectral and spatial features, respectively, to construct multi-feature kernels. The training sets are randomly selected from the available labeled samples and that the remaining samples are used for validation. We optimized the SVM parameters using tenfold cross-validation.

After the spectral and the spatial characterization, each pixel has been presented by five vectors:

- $x^{PCA}$ is a spectral vectors that contains the first five principals components.
- $x^{ICA}$ is a spectral vector containing the six independents components.
-$x^{Neigh}$ is a spatial vector that contains the average of neighborhood pixels in a window of size 3×3.
- $x^{EMAP}$ a spatial vector that contains the EMAP of each pixel. EMAP were built according to the used attributes and thresholds presented in [25]: threshold values in the range of 2,5% - 10% with a step of 2,5% for the standard deviation attribute and thresholds of 200, 500 and 1000 for the area attribute.
- $x^{Text}$ is a spatial vector that contains the textural features computed from the first three principals components. Note that we applied tree sliding windows: 3×3, 9×9 and 15×15 and we used four directions to calculate these features (LM, VAR, ASM and ENT).

Must indicate that the number of the used principals components in $x^{PCA}$ and the number of the adopted independents components for $x^{ICA}$ have been experimentally fixed according to the spectral classification of the image when using 10 labeled samples in each class (Figure 4).




(b)

Figure 4. OA variation according to the variation of the number of principal components (a) and the number of independent components (b).

In order to analyze the impact of the multi-feature kernel and to select the most suitable for the classification of the "Indian Pines" data set, we analyze the performance of the proposed method for the tree kernels: $K^{PCA+ICA+Neigh+Text+EMAP}$, $\mu K^{PCA+ICA+Neigh+Text+EMAP}$ and $\alpha K^{PCA+ICA+Neigh+Text+EMAP}$ in a classification with generating in each class 10 samples about 10 available training samples. Figure 5 shows the OAs and kappa obtained by the proposed classification algorithm according to the applied kernels. Note that we used $\mu$=0.25 for $\mu K^{PCA+ICA+Neigh+Text+EMAP}$ and (0.1,0.1,0.2,0.2,0.4) for $\alpha K^{PCA+ICA+Neigh+Text+EMAP}$. This chose has been fixed experimentally (Figure 6). As shown in Figure 3, the performance of the proposed classification algorithm increases when using $\alpha K^{PCA+ICA+Neigh+Text+EMAP}$ which valorize the spatial features extracted by EMAP ($\alpha_5$=0.4 > $\alpha_3$=$\alpha_4$=0.2 > $\alpha_1$= $\alpha_2$=0.1), it yielded much better OA and kappa results (OA= 82% and kappa=81,73%). Furthermore, we can note that the weighted summation kernel introducing a trade-off ($\mu$) between spectral and spatial kernels with $\mu$=0.25 performs more accurately than the direct summation kernel.

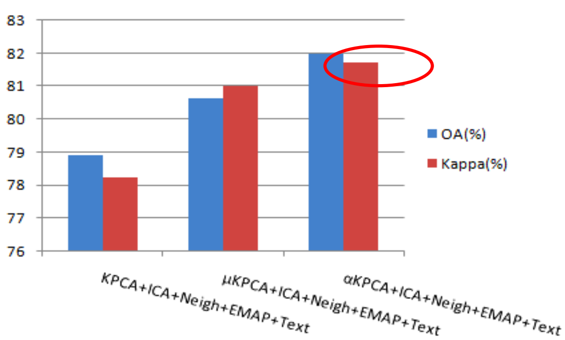
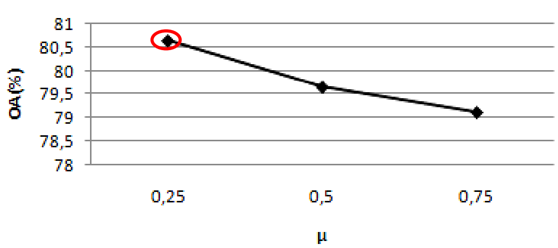
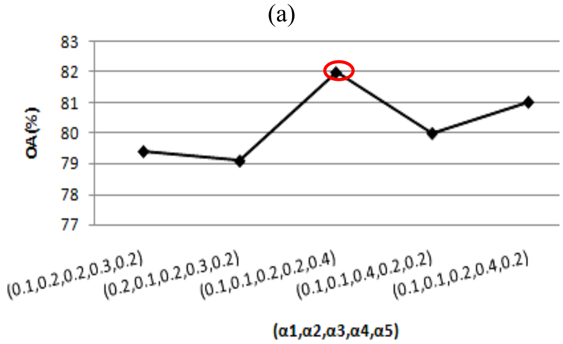



Figure 5: Resulted OA and Kappa coefficient according to the adopted multi-feature kernel




(a)




(b)

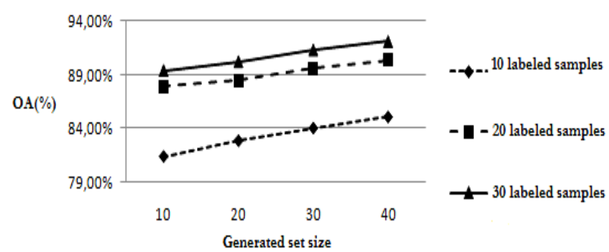
Figure 6. Variation of OA according to the variation of $\mu$ for the kernel $\mu K^{PCA+ICA+Neigh+Text+EMAP}$ (a) and ($\alpha_1$,$\alpha_2$,$\alpha_3$,$\alpha_4$,$\alpha_5$) for the kernel $\alpha K^{PCA+ICA+Neigh+Text+EMAP}$(b).

To show the advantage of the oversampling, we note in Table I the classification results obtained for different number of training samples after oversampling and without oversampling. In this experiment, we used cubic spline interpolation for the spectral and the spatial features to create new labeled samples and we adopted $\alpha K^{PCA+ICA+Neigh+EMAP+Text}$ as a multi-feature kernel.

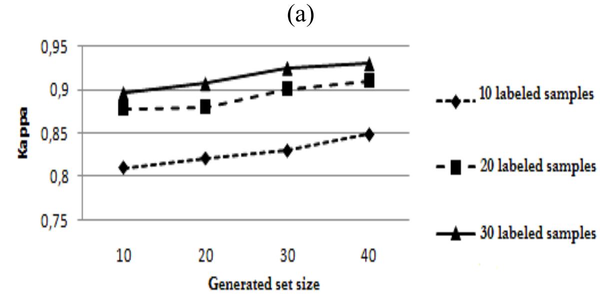
Table I illustrate the average of the OA followed by the standard deviation (±) and the kappa coefficient obtained after ten Monte Carlo runs. By adopting oversampling, the proposed method significantly improved the classification results obtained by the considered classification without oversampling for all the adopted size of training set (10, 20 and 30). For instance, the generation of 40 samples about 10 labeled examples obtained an OA of 85.09%, 6.09% larger than that obtained by SVM without oversampling. As a result, the obtained samples after oversampling improves the accuracy of the supervised classifier (SVMs with multi-feature kernel). Notice also that the increase in the number of generated data improves significantly the performance of the classification (Figure 7) which indicates the advantage these samples that increase the ability of SVM to find the optimal separator hyperplan.

Table 1: OA and kappa coefficient (in parenthesis) obtained for the Indian Pines data set

| Labeled samples | Number of generated samples | | | | |
|---|---|---|---|---|---|
| | 0 | 10 | 20 | 30 | 40 |
| 10 | 79% ± 1.3 (0.789) | 81.32% ± 1.25 (0.81) | 82.5% ± 1.4 (0.821) | 84.04% ± 0.54 (0.83) | 85.09% ± 0.64 (0.849) |
| 20 | 85.56 ± 1.2 (0.86) | 87.86 ± 1.11 (0.878) | 88.39% ± 0.66 (0.89) | 89.54% ± 0.67 (0.9) | 90.34% ± 1.1 (0.91) |
| 30 | 87.68 ± 0.92 (0.85) | 89.34% ± 1.07 (0.896) | 90.19% ± 1.15 (0.907) | 91.3% ± 1.1 (0.92) | 92.11% ± 0.56 (0.93) |




(a)




(b)

Figure 7: Variation of OA (a) and kappa coefficient (b) according to the increase in generated set size.

Focusing on the oversampling step, we have indicated in the description of the proposed classification method that we have implemented tree interpolation techniques to generate new training data: linear interpolation, cubic spline interpolation and Lagrange interpolation. Then to analyze the impact of these techniques and to select the most adequate for each feature, we note in table 2 the spectral, the spatial and the spectral-spatial classification results obtained after the generation of 20 data from 10 labeled samples by applying different interpolation methods. Table 2, illustrates for each interpolation method the OA, AA, kappa statistic coefficient (k), and individual class accuracy (in percent) results achieved by the spectral classification ( two spectral classifications: using characteristics extracted by PCA and characteristics extracted by ICA), the spatial classification (tree spatial classifications: using characteristics extracted by the average of neighborhoof pixels, textural features and characteristics extracted by EMAP) and the spectral-spatial classification assured by the proposed method. It is remarkable that the uses of Lagrange interpolation to create new spectral features (PCA, ICA) and spatial features extracted by EMAP and the average of neighborhood pixels leads to have more accurate spectral and spatial classifications than that resulted after the application of the other methods: cubic spline and linear interpolation. For textural features, it's clear that the linear interpolation provided the highest accuracy. Then, for the spectral-spatial classification we applied the Lagrange interpolation for feature extracted by PCA, ICA, Neigh, and EMAP and the linear interpolation for the textural features. This combination provided high performance, it obtained an OA of 83.87% and kappa coefficient of 0,8385, 1.37% and 0.017 larger than these obtained when we used cubic spline interpolation for all features (indicated in table 1). This indicate that the samples generated by this combination are properly created refer to their similarity to the oversampled data in each class.

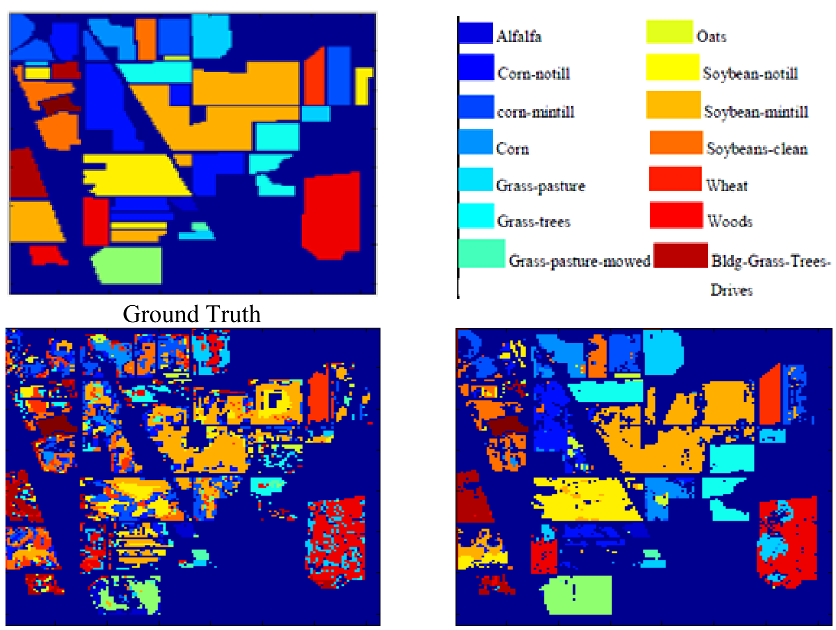
Figure 8 shows the ground truth and the classification result obtained without-oversampling and by the proposed method for the AVIRIS Indian Pines scene. The advantage of the proposed classification approach is clearly appreciable in this figure.

Table 2: Table 2: OA, AA an kappa coefficient obtained for the AVIRIS Indian Pines data set

| Class | Test samples | Classification with oversampling | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Spectral classification | | | | | | Spatial classification | | |
| | | ACP | | | ACI | | | EMAP | | |
| | | Lin | Spl | Lag | Lin | Spl | Lag | Lin | Spl | Lag |
| Alfalfa | 44 | 77,27 | 79,55 | 79,55 | 86,36 | 86,36 | 79,55 | 86,36 | 86,36 | 86,36 |
| Corn-notill | 1424 | 22,33 | 8,64 | 15,45 | 27,46 | 21,35 | 23,88 | 30,34 | 30,34 | 30,68 |
| Corn-mintill | 824 | 22,21 | 21 | 38,47 | 34,10 | 28,40 | 43,45 | 34,10 | 27,55 | 27,55 |
| Corn | 224 | 72,32 | 70,54 | 82,59 | 68,30 | 65,18 | 73,21 | 41,07 | 49,55 | 49,55 |
| Grass/pasture | 487 | 68,38 | 65,50 | 58,11 | 84,80 | 82,75 | 75,15 | 82,75 | 82,75 | 82,75 |
| Grass/trees | 737 | 81,68 | 84,12 | 86,43 | 83,18 | 88,60 | 87,38 | 87,89 | 87,89 | 85,89 |
| Grass/pasture-mowed | 16 | 81,25 | 87,50 | 87,50 | 68,75 | 87,50 | 87,50 | 93,75 | 93,75 | 93,75 |
| Hay-windrowed | 479 | 88,10 | 89,14 | 81 | 66,39 | 68,89 | 89,77 | 91,65 | 91,65 | 91,65 |
| Oats | 10 | 90 | 90 | 100 | 80 | 90 | 100 | 100 | 90 | 100 |
| Soybean-no till | 958 | 43,63 | 46,56 | 34,76 | 15,76 | 48,33 | 13,47 | 48,15 | 48,15 | 55,5 |
| Soybean-min till | 2458 | 51,51 | 31,08 | 44,26 | 48,41 | 8,01 | 57,28 | 48,70 | 48,70 | 48,70 |
| Soybean-clean till | 604 | 43,21 | 36,26 | 44,87 | 45,20 | 26,16 | 64,57 | 66,23 | 66,23 | 66,23 |
| Wheat | 202 | 63,37 | 74,75 | 79,70 | 52,97 | 56,93 | 68,81 | 95,05 | 95,05 | 95,05 |
| Woods | 1284 | 28,27 | 47,51 | 47,90 | 47,66 | 52,65 | 58,18 | 75,86 | 75,86 | 75,86 |
| Bldg-Grass-Trees-Drives | 370 | 40 | 62,16 | 50,81 | 18,92 | 26,22 | 20,54 | 62,16 | 71,35 | 71,35 |
| Stone-Steel-Towers | 85 | 91,76 | 91,76 | 97,65 | 90,59 | 91,76 | 91,76 | 91,76 | 97,65 | 97,65 |
| OA (%) | | 69,59 | 68,84 | **70,59** | 68,43 | 66,61 | **71,44** | 72,11 | 72,98 | **74,02** |
| AA (%) | | 61,63 | 61,63 | **64,32** | 57,43 | 58,07 | **64,66** | 69,82 | 70,67 | **71,15** |
| kappa | | 0,6026 | 0,5872 | **0,6089** | 0,6152 | 0, 6181 | **0.67** | 0,721 | 0,721 | **0,721** |

Table 2 (suite)

| **Class** | **Test samples** | **Classification with oversampling** | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **Spatial classification** | | | | | | **Multi-feature classification** | |
| | | **Neighborhood** | | | **Textural features** | | | | |
| | | **Lin** | **Spl** | **Lag** | **Lin** | **Spl** | **Lag** | **Without oversampling** | **Lag_Lag_Lag_Lag_Lin** |
| Alfalfa | 44 | 88,64 | 86,36 | 86,36 | 84,09 | 88,64 | 84,09 | 90,91 | 93,18 |
| Corn-notill | 1424 | 19,87 | 21,14 | 22,68 | 9,27 | 13,41 | 10,60 | 30,34 | 38,97 |
| Corn-mintill | 824 | 28,28 | 31,92 | 27,55 | 16,38 | 17,23 | 24,51 | 34,59 | 57,28 |
| Corn | 224 | 45,54 | 41,52 | 49,55 | 41,07 | 41,07 | 37,05 | 82,59 | 87,05 |
| Grass/pasture | 487 | 47,23 | 45,38 | 60,57 | 44,35 | 40,66 | 35,93 | 74,33 | 82,96 |
| Grass/trees | 737 | 82,77 | 81,82 | 85,89 | 31,34 | 31,48 | 27,68 | 89,69 | 93,49 |
| Grass/pasture-mowed | 16 | 93,75 | 93,75 | 93,75 | 81,25 | 87,50 | 87,50 | 87,50 | 87,50 |
| Hay-windrowed | 479 | 94,36 | 93,53 | 91,65 | 59,71 | 60,54 | 55,74 | 92,28 | 95,82 |
| Oats | 10 | 100 | 100 | 100 | 80 | 80 | 90 | 100 | 100 |
| Soybean-no till | 958 | 49,16 | 49,69 | 49,58 | 27,56 | 34,66 | 43,01 | 55,22 | 68,06 |
| Soybean-min till | 2458 | 48,58 | 48,17 | 48,70 | 45,77 | 45,36 | 42,35 | 52,73 | 59,72 |
| Soybean-clean till | 604 | 70,03 | 69,04 | 66,23 | 40,89 | 30,79 | 39,57 | 50,83 | 62,42 |
| Wheat | 202 | 95,05 | 95,54 | 95,05 | 85,15 | 90,10 | 86,63 | 98,51 | 99,01 |
| Woods | 1284 | 63,16 | 72,98 | 75,86 | 69,55 | 63,24 | 64,49 | 62,31 | 62,69 |
| Bldg-Grass-Trees-Drives | 370 | 70,27 | 71,62 | 71,35 | 73,78 | 63,24 | 54,59 | 45,68 | 65,95 |
| Stone-Steel-Towers | 85 | 100 | 100 | 97,65 | 98,82 | 98,82 | 100 | 98,82 | 98,82 |
| OA (%) | | 70,83 | 71,5 | **73,12** | **69,02** | 65,16 | 67,62 | 78,79 | **83,87** |
| AA (%) | | 68,54 | 68,91 | **70,15** | **55,56** | 55,42 | 55,23 | 71,64 | 78,31 |
| kappa | | 0,574 | 0,5944 | **0,6187** | **0,475** | 0,4657 | 0,4662 | 0,78 | 0,8385 |




Without oversampling 78,79% With oversampling 82,87%

Figure 8: Classification maps obtained for the AVIRIS Indian Pines scene

## 4. Conclusion

In this paper, we have presented a new spectral-spatial classification approach which combines various spectral and spatial features via multi-feature kernels and generate new training samples to solve two problems widely proposed in the classification of hyperspectral images which are the difficulty of the choice of the applied characterization methods and the availability of a limited number of labeled samples. It investigates the oversampling based on interpolation techniques to increase the size of training set. By using the kernel $\alpha K^{PCA+ICA+Neigh+Text+EMAP}$ which valorize the spatial features computed by EMAP and by adopting the Lagrange interpolation for features extracted by PCA, ICA, the average of neighborhood pixels and EMAP and the linear interpolation for textural features in the oversampling step, the proposed method provides good accuracies when compared with the spectral classification, the spatial classification and the classification without oversampling. Combining multi-feature kernels and oversampling provides competitive and encouraging results. Further work should be focused on the exploitation of active learning algorithms to improve the quality of the generated samples in the oversampling step.

## 5. References

[1] F. M. Lacar, M. M. Lewis, and I. T. Grierson, "Use of hyperspectral imagery for mapping grape varieties in the Barossa Valley, South Australia," in IEEE International Geoscience and Remote Sensing Symposium 2001 (IGARSS '01), vol.6, pp. 2875–2877, 2001.

[2] E. K. Hege et al. "Hyperspectral imaging for astronomy and space surveillance," in Proc. SPIE 5159, Imaging Spectrometry IX, 380, 2004.

[3] P.W. Yuen and M. Richardson, "An introduction to hyperspectral imaging and its application for security, surveillance and target acquisition," The Imaging Science Journal, vol. 58, no. 5, pp. 241–253, 2010.

[4] J. B. Dias et al., "Hyperspectral remote sensing data analysis and future challenges," IEEE Geoscience and Remote Sensing Magazine, vol. 1, no. 2, pp. 6–36, 2013.

[5] H. Zhang, J. Li, Y. Huang, and L. Zhang, "A nonlocal weighted joint sparse representation classification method for hyperspectral imagery," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 6, pp. 2056–2065, Jun. 2014.

[6] H. Yuan, Y. Yan Tang, Y. Lu, L. Yang, and H. Luo, "Hyperspectral image classification based on regularized sparse representation," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 7, no. 6, pp. 2174–2182, Jun. 2014.

[7] G. Camps-Valls,, L. Bruzzone, " kernel methods for remote Sensing data analyzes", J.wiley and Sons, NJ, USA, 2009.

[8] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla "Composite Kernels for Hyperspectral Image Classification", IEEE geoscience and remote sensing letters, vol.3, no.1, pp. 93 - 97, Jan. 2006.

[9] M. Fauvel, J. A. Benediktsson, J. chanussot and J.R. Sveinsson " Spectral and Spatial Classification of Hyperspectral Data Using SVMs and Morphological Profiles", IEEE transaction on geoscience and remote sensing, vol.46, no.11, pp. 3804 - 3814, Nov. 2008.

[10] X. Kang, S. Li and J. A. Benediktsson "Spectral-spatial Hyperspectral Image Classification with Edge-preserving Filtering", IEEE transactions on geoscience and remote sensing, vol. 52 , no. 5, pp. 2666 - 2677, May 2014.

[11] X. Huang and L. Zhang "An SVM Ensemble Approach Combining Spectral, Structural, and Semantic Features for the Classification of High-Resolution Remotely Sensed Imagery", IEEE transactions on geoscience and remote sensing, vol. 51, no. 1, pp. 257 - 272, Jan. 2013.

[12] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias and J. A. Benediktsson " Generalized Composite Kernel Framework for Hyperspectral Image Classification", IEEE transactions on geoscience and remote sensing, vol. 51, no. 9, pp. 4816 - 4829, Sep. 2013.

[13] P. Ghamisi, M. S. Couceiro, M. Fauvel and J. A. Benediktsson "Integration of Segmentation Techniques for Classification of Hyperspectral Images", IEEE on geoscience and remote sensing letters, vol. 11, no. 1, pp. 342-346, JANUARY 2014.

[14] R. Ben Salem, K. S. Ettabaa and M. A. Hamdi " Spectral-spatial classification of hyperspectral images using different spatial features and composite kernels", in IEEE International Image Processing, Applications and Systems Conference 2014 (IPAS'14), pp. 1- 7, Nov.2014.

[15] Rodarmel and J. Shan " Principal Component Analysis for Hyperspectral Image Classification", Surveying and Land Information Systems, vol. 62, no. 2, pp.115-123 , 2002.

[16] Q. Du "Independent component analysis to hyperspectral image classification", Proc. SPIE 5546, Imaging Spectrometry X, 366, Oct.2004.

[17] M.D. Mura, J.A.Benediktsson, B; Waske, L. Bruzzone, "Extended profiles with morphological attribute filters for the analyses of hyperspectral data," Int. J. Remote sens., vol. 31, no. 22, pp.5975-5991, Jul. 2010

[18] X. Huanga, L. Zhangaand P. Lia " A multiscale feature fusion approach for classification of very high resolution satellite imagery based on wavelet transform", International Journal of Remote Sensing, vol. 29, no. 20, pp. 5923-5941, Oct. 2008.

[19] X. Huang, L. Zhang and P. Li" An Adaptive Multiscale Information Fusion Approach for Feature Extraction and Classification of IKONOS Multispectral Imagery Over Urban Areas", ", IEEE on geoscience and remote sensing letters, vol. 4, no. 4, pp. 654 - 658, Oct. 2007.

[20] S. Serpico and G. Moser, "Extraction of spectral channels from hyperspectral images for

classification purposes," IEEE Trans. Geosci. Remote Sens., vol. 45, no. 2, pp. 484–495, Feb. 2007.

[21] R. Ben Salem, K. Saheb Ettabaa and M. A. Hamdi" Supervised spectral-spatial hyperspectral image classification based on oversampling and composite kernels", ICGST-GVIP, vol.15, no.2, pp.27-39, Dec. 2015.

[22] M. Volpi, D. Tuia, F. Bovolo, M. Kanevski, L. Bruzzone " Supervised change detection in VHR images using contextual information and support vector machines", International Journal of Applied Earth Observation and Geoinformation, vol. 20, pp. 77–85, 2013.

[23] M.D. Mura, J.A.Benediktsson, B. Waske, L. Bruzzone, "Extended profiles with morphological attribute filters for the analyses of hyperspectral data," Int. J. Remote sens., vol. 31, N. 22, pp.5975-5991, Jul. 2010.

[24] G. Mercier, M. Lennon, " Support vector machines for hyperspectral image classification with spectral based kernels," International symposium on geoscience and remote sensing (IGARSS 2003), vol.1, pp. 288 - 290 , July 2003.

[25] J.A. Benediktsson, J.A. Palmason et J.R. Sveinsson "Classification of hyperspectral data from urban areas based on extended morphological profiles", IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 3, pp. 480-491, Mar. 2005.

## Biographies

Rafika Ben Salem received the Eng. and M. degrees from National school of engineering, Sousse at 2009 and 2011. Currently she is a PhD student at National institute of applied sciences and technology, Tunis in the MMA Laboratory. Her research interests include remote sensing image interpretation and classification.

Karim Saheb Ettabaa received the M.E. and Dr. Eng. degrees from ENSI, Manouba, Tunisia, in 2004 and 2007, respectively. He is a Permanent Researcher at laboratory RIADI, National School of Computer Sciences Engineering and ITI, Telecom Bretagne, Brest Iroise Technopole CS 81828. His research interest includes image processing, data mining, artificial intelligence, pat- tern recognition, and their application to remote sensing

Mohamed Ali HAMDI received the Ph.D. degree in image processing, from the National University of engineering of TUNIS, He is Assistant Professor of Applied and Computational Mathematics and electronic, National institute of applied sciences and technology. His research interests are in the areas of digital signal processing (DSP), statistical estimation and their applications to signal and image processing and scientific computing.